2006

Microarray Based Comparison of Two *Escherichia coli* O157:H7 Lineages

Scot E. Dowd
USDA-ARS

Hiroshi Ishizaki
Department of Grazing Animal Production, National Institute of Livestock and Grassland Science, Nasushiobara, Japan

BMC Microbiology

BioMed Central

Research article

Microarray based comparison of two *Escherichia coli* O157:H7 lineages

Scot E Dowd*[1] and Hiroshi Ishizaki[2]

Address: [1]Livestock Issues Research Unit, USDA-ARS, Lubbock, TX, USA and [2]Department of Grazing Animal Production, National Institute of Livestock and Grassland Science, Nasushiobara, Tochigi 329-2793, Japan

Email: Scot E Dowd* - sdowd@lbk.ars.usda.gov; Hiroshi Ishizaki - hishizak@affrc.go.jp

* Corresponding author

Published: 15 March 2006

*BMC Microbiology*2006, **6**:30 doi:10.1186/1471-2180-6-30

Received: 04 January 2006
Accepted: 15 March 2006

This article is available from: http://www.biomedcentral.com/1471-2180/6/30

Abstract

Background: Previous research has identified the potential for the existence of two separate lineages of *Escherichia coli* O157:H7. Clinical isolates tended to cluster primarily within one of these two lineages. To determine if there are virulence related genes differentially expressed between the two lineages we chose to utilize microarray technology to perform an initial screening.

Results: Using a 610 gene microarray, designed against the *E. coli* O157 EDL 933 transcriptome, targeting primarily virulence systems, we chose 3 representative Lineage I isolates (LI groups mostly clinical isolates) and 3 representative Lineage II isolates (LII groups mostly bovine isolates). Using standard dye swap experimental designs, statistically different expression ($P < 0.05$) of 73 genes between the two lineages was revealed. Result highlights indicate that under *in vitro* anaerobic growth conditions, there is up-regulation of *stx2b*, *ure*D, curli (*csg*AFEG), and stress related genes (*hsl*J, *csp*G, *ibp*B, *ibp*A) in Lineage I, which may contribute to enhanced virulence or transmission potential. Lineage II exhibits significant up-regulation of type III secretion apparatus, LPS, and flagella related transcripts.

Conclusion: These results give insight into comparative regulation of virulence genes as well as providing directions for future research. Ultimately, evaluating the expression of key virulence factors among different *E. coli* O157 isolates has inherent value and the interpretation of such expression data will continue to evolve as our understanding of virulence, pathogenesis and transmission improves.

Background

Kim et al., [1] utilized octamer-based genome scanning to evaluate genome diversity among *E. coli* O157 isolates. Based upon this genetic fingerprinting method they noted two distinct lineages of this pathogen, one of which tended to cluster the majority of human isolates utilized in their study, and the second which grouped together isolates primarily of bovine origin. They suggested that one of these lineages (Lineage II) may not efficiently transmit to humans from bovine sources. Pradel et al. [2] also found that there were distinct lineages among isolates derived from patients with hemolytic-uremic syndrome (HUS) when evaluated genetically using a combination of *stx2*-RFLP (restriction fragment length polymorphism analyses), *stx2* variant, and plasmid profile analyses. They also suggested that there may be a separate lineage, which was more virulent for humans, along with a lineage, which may not be as pathogenic. Yang et al. [3] utilized a

Table 1: This table presents the genes that were consistently and statistically up regulated in the hypothetically more pathogenic lineage I strains.

Name	average mean log ratio (635/532)	average fold increase	One Sample t-Test (p)	definition	accession
hslJ	-2.04	4.10	0.017036	heat shock protein hslJ	NP_287767.1
cspG	-1.97	3.91	0.018217	homolog of Salmonella cold shock protein	NP_286926.1
ibpB	-1.82	3.54	0.037563	heat shock protein	NP_290325.1
ibpA	-1.59	3.02	0.031977	heat shock protein	NP_290324.1
Fimb 10	-1.58	2.98	0.011803	putative fimbrial protein	NP_290361.1
trxC	-1.39	2.63	0.018467	putative thioredoxin-like protein	NP_289141.1
feoB	-1.31	2.47	0.004959	ferrous iron transport protein B	NP_289949.1
ureD	-1.28	2.43	0.038695	putative urease accessory protein D	NP_287085.1
chap3	-1.26	2.39	0.030462	possible chaperone	NP_288034.1
ydeA	-1.23	2.35	0.026584	putative resistance/regulatory protein	NP_287624.1
ureD_2	-1.23	2.34	0.050155	putative urease accessory protein D	NP_287085.1
secret3	-1.09	2.13	0.024694	putative secreted protein	NP_288394.1
fliC	-1.06	2.08	0.018058	flagellar biosynthesis; flagellin, filament structural protein	NP_288384.1
stx2A	-1.00	2.01	0.021712	shiga-like toxin II A subunit encoded by bacteriophage BP-933W	NP_286976.1
terW	-0.95	1.93	0.044205	unknown associated with putative tellurite resistance	NP_286699.1
rfaH	-0.93	1.90	0.007198	transcriptional activator affecting biosynthesis of lipopolysaccharide core, F pilin, and haemolysin	NP_290472.1
phoB	-0.87	1.82	0.006296	positive response regulator for pho regulon, sensor is PhoR (or CreC)	NP_286137.1
ureA_2	-0.81	1.75	0.050244	putative urease structural subunit A (gamma)	NP_287086.1
helicase1	-0.71	1.64	0.01929	putative ATP-dependent helicase	NP_288767.1
yegW	-0.71	1.63	0.015902	putative transcriptional regulator	NP_288603.1
ureB_2	-0.70	1.62	0.038644	putative urease structural subunit B (beta)	NP_287087.1
typeIIIap7	-0.66	1.58	0.033765	type III secretion apparatus protein	NP_289424.1
ycjZ	-0.61	1.53	0.038253	putative transcriptional regulator LYSR-type	NP_287857.1
ykgA	-0.60	1.51	0.000965	putative AraC-like transcriptional regulator	NP_286025.1
sbmA	-0.59	1.50	0.050279	sensitivity to microcin B17, possibly envelop protein	NP_286115.1
YjhS	-0.57	1.49	0.03404	orf, hypothetical protein	NP_290925.1
ybbK	-0.53	1.45	0.037791	putative protease	NP_286238.1
virulense 3	-0.50	1.42	0.038974	putative virulence gene	NP_290837.1
pldA	-0.49	1.41	0.031981	outer membrane phospholipase A	NP_290453.1
yejH	-0.48	1.40	0.008115	putative ATP-dependent helicase	NP_288767.1
transreg	-0.34	1.27	0.033811	putative membrane spanning transport protein	NP_286232.1
hydH	-0.27	1.21	0.00039	sensor kinase for HydG, hydrogenase 3 activity	NP_290635.1

lineage-specific polymorphism assay consisting of 6 genetic markers and found that they could differentiate two lineages of *E. coli* O157 indicating that the occurrence of these two lineages may be widespread. Barkocy-Gallagher [4] using *Xba1* RFLP analysis also found distinct clusters of *E. coli* O157, including a cluster where most isolates lacked flagella and *stx1* genes, leading them to suggest the potential for the existence of clustered isolates having differential abilities to cause disease.

The expression of several virulence factors in relation to the existence of two lineages of EHECs have been evaluated as well. McNally et al. [5] found clear differences in the expression of locus of enterocyte effacement (LEE)-encoded factors between different strains. It was found that, *EspD*, when used as an indicator of LEE expression, was expressed at higher concentrations in the majority of strains that were of human origin (15 of 20) compared with only a few (4 of 20) isolates that were of bovine origin ($P < 0.001$). They concluded that a subset of *E. coli* O157 isolates (stx^+ eae^+) in cattle were capable of causing severe disease in humans. Another study evaluating gene expression conducted by Richie et al., [6] found that HUS derived isolates expressed higher concentrations of *stx2* than bovine derived isolates.

Based upon the proposed existence of a less pathogenic lineage of *E. coli* O157, it has been postulated that much of the Class I recall of millions of pounds of meat annually [7] might be greatly reduced. However, even if a separate lineage of *E. coli* O157 (conclusively proven not to cause disease in humans) were identified and concrete

methods for differentiating this lineage developed, it would still be unlikely (because of liability issues) to have the suggested impact on the meat industry. Yet, the study of genetic differences between two lineages of this pathogen that possess different virulence or transmission potential could still have wide ranging and significant economic or scientific benefits. For example, if a specific lineage could be more readily eradicated during the farm to fork process, based upon their genetic differences, this might indirectly have the originally intended effect of reducing the volume of Class I recalls. In addition, from a purely scientific standpoint, clues as to why certain isolates may be more pathogenic or more easily transmitted, based upon genetic differences, is of obvious importance in the study of virulence.

Results and discussion

Microarray analyses, validated by quantitative PCR, showed that, of the 610 genes on the array, 179 genes were consistently and differentially regulated between the two lineages. Of these 179 regulated genes, 73 transcripts showed statistically significant ($p < 0.05$) differences in expression of greater than 1.2 fold (Table 1 and Table 2) between each member of the two lineages. Table 1 shows those transcripts whose expression was greater ($P < 0.05$) in each of the LI isolates. Three heat shock and one cold shock protein transcripts were the most upregulated in the LI isolates compared to the LII. In LII isolates cyoE, hscA, and fimbrial subunit 1 were most highly upregulated compared to LI. Table 2 shows those transcripts whose expression was statistically higher ($P < 0.05$) in each of the LII isolates. Six transcripts that exhibited enough expression difference to be evaluated by CT using quantitative PCR were chosen at random from these 73 and Q-PCR performed as a validation method. These included *ure*D, cyoE, hscA, nrfB [see Additional file 1], chap4, and *stx*2B. Results of Q-PCR were found to agree in each instance with the results of the microarray experiment. Supplementary dataset 1 [see Additional file 1] provides a listing of the 106 genes that were shown to be consistently up-regulated or down-regulated as part of the microarray experiment, but which did not fully meet the stringent selection and statistical requirements additional supplemental dataset 2 provides all the genes on the array.

Results of the microarray experiments showed that the LI isolates express higher transcription of ureD (Table 2), as well as ureA, ureB, ureC (supplemental data), compared to LII. In addition stx2B (Table 2) and stx2A (supplemental data) transcripts are detected in higher abundance in Lineage 1. Lineage I also exhibits up-regulation of key fimbria related transcripts, especially fliC, fliT, and fliP. Other attachment related transcripts csgA, csgF, csgE, and csgG (curli) were also up-regulated, which could also be highly significant in promotion of pathogenesis [8-12]. When using all of the regulated genes as a single data set for Gene Ontology [13] based analyses, it was found that, up-regulation of genes associated with regulation of urease activity, GTP binding, metabolism, nitrogen metabolism and regulation of transcription were statistically ($p < 0.05$) more represented in LI isolates (Table 3). In LII isolates peptidase activity, transferase activity, and DNA binding activity were statistically more represented ($p < 0.05$). These differences could point to a fundamental difference in the environmental response and control networks of these lineages that promotes survival and differential expression of virulence attributes in response to specific environments and hosts. These types of control networks could be the key to understanding differential virulence or transmission potential if such a phenomenon could be proven to exist within the O157 serogroup.

Stx2

The role of *stx2* in pathogenesis is well accepted [14-18] and up-regulation of constitutive stx2 expression in the hypothesized more pathogenic LI isolates may not be a surprising finding. The up-regulation of *stx2B* and *stx2A* [see Additional file 1] transcripts is accompanied by up-regulation of regulatory genes associated with Stx2 expression. A complicated network of interactions between the o*raA* (*recX*), *dinI*, *lexA*, *umuD*, SSB, *recA*, *psiB* and possibly other unidentified proteins, act in the regulation of RecA function. The role of *recA* as part of an SOS response is to cleave repressors that in addition to the SOS response ultimately lead to Stx2 production [19]. *OraA* (also known as *recX*) and *dinl* are coregulators (competing regulators) of *recA* [20] and both were up-regulated in LI isolates along with the *stx2* subunit transcripts. *OraA* is thought to be co-transcribed with *recA* during SOS response [21]. *RecA* specific oligos were not included in the array but we might expect that being co-transcribed along with *oraA* that it would likely be up-regulated in LI as well. *PsiB*, (supplemental) is also up-regulated in LI and thought to prevent ssDNA from inducing an SOS response by inhibiting activation of *recA* protein [22]. *PsiB*, is found on many conjugative plasmids near the origin of conjugative transfer and has anti-recombinase activities [23]. Expression of the *dinI* protein of *E. coli* inhibits both the co-protease and recombinase activities of *recA* in vivo [24]. Yet, in spite of all of the regulators of SOS response in LI isolates, we still observe a significant up-regulation of *stx2a* and *stx2b* transcripts which have been shown to be expressed as part of an SOS response [25-27].

With up-regulation of *dinl*, *psiB*, *oraA* and also with the up-regulation of *stx2a* and *stx2b* and various other genes related to stress response it could be an indication that LI isolates do have differentially regulated pathways the enhance its toxin expression potential. It does appear that the current LI isolates have a modified regulatory system

Table 2: This table presents genes that were consistently and statistically up regulated in the hypothetically less pathogenic lineage II strains

Name	average mean log ratio (635/532)	average fold increase	One Sample t-Test (p)	definition	accession
inaA	0.57	1.49	0.028507	pH-inducible protein involved in stress response	NP_288811.1
nupG	0.60	1.51	0.002513	transport of nucleosides, permease protein	NP_289536.1
ompR	0.61	1.53	0.019932	response regulator (sensor, EnvZ) affecting transcription of ompC and ompF: outer membrane protein synthesis	NP_289945.1
flgC	0.62	1.54	0.008244	flagellar biosynthesis, cell-proximal portion of basal-body rod	NP_287208.1
secD	0.64	1.55	0.050195	protein secretion; membrane protein, part of the channel	NP_286147.1
recB	0.75	1.68	0.050257	DNA helicase, ATP-dependent dsDNA/ssDNA exonuclease V subunit, ssDNA endonuclease	NP_289372.1
hemY	0.78	1.72	0.045026	a late step of protoheme IX synthesis	NP_290430.1
fhuA	0.79	1.73	0.015243	outer membrane protein receptor for ferrichrome, colicin M, and phages T1, T5, and phi80	NP_285846.1
cheB	0.81	1.75	0.014779	response regulator for chemotaxis (cheA sensor); protein methylesterase	NP_288320.1
cutC	0.83	1.77	0.017209	Copper homeostasis protein	NP_288311.1
secF	0.96	1.94	0.033545	protein secretion, membrane protein	NP_286148.1
wecF	0.97	1.95	0.021373	TDP-Fuc4NAc:lipidII transferase; synthesis of enterobacterial common antigen (ECA)	NP_290425.1
frdD	0.97	1.95	0.042158	fumarate reductase, anaerobic, membrane anchor polypeptide	NP_290786.1
espP	1.07	2.10	0.023407	EspP	NP_052685.1
kfras	1.08	2.11	0.040371	KfraS	NP_052633.1
etpJ	1.10	2.15	0.016101	EptJ	NP_052615.1
toxB	1.11	2.15	0.020589	toxin B	NP_052665.1
fliY	1.11	2.16	0.001708	putative periplasmic binding transport protein	NP_288381.1
wzx	1.22	2.33	0.014144	O antigen flippase Wzx	NP_288543.1
etpH	1.28	2.42	0.009559	EtpH	NP_052613.1
etpO	1.28	2.43	0.021162	EptO	NP_052620.1
fumC	1.31	2.48	0.01514	a late step of protoheme IX synthesis	NP_290430.1
ydeW	1.34	2.53	0.001173	putative transcriptional regulator, sorC family	NP_287642.1
etpI	1.39	2.62	0.005332	EtpI	NP_052614.1
etpM	1.57	2.98	0.000193	EtpM	NP_052618.1
cspD	1.59	3.01	0.051261	cold shock protein	NP_286652.1
oppC	1.60	3.03	0.009279	homolog of Salmonella oligopeptide transport permease protein	NP_287488.1
usher2	1.66	3.15	0.02289	putative fimbrial usher protein	NP_287650.1
cspC	1.67	3.19	0.048484	cold shock protein	NP_288259.1
etpN	1.68	3.20	0.015507	EptN	NP_052619.1
etpK	1.70	3.25	0.000151	EtpK	NP_052616.1
rpoB	1.74	3.34	0.005075	RNA polymerase, beta prime subunit	NP_290619.1
argT	1.81	3.51	0.007679	lysine-, arginine-, ornithine-binding periplasmic protein	NP_288884.1
fimbera l4	1.84	3.58	0.0179	putative fimbrial protein	NP_290128.1
chaper 2	1.86	3.64	0.002084	putative fimbrial chaperone	NP_287649.1
fumA	1.88	3.68	0.005453	fumarase C= fumarate hydratase Class II; isozyme	NP_288046.1
etpL	1.97	3.92	0.009254	EtpL	NP_052617.1
rpoC	2.01	4.04	0.021043	RNA polymerase, beta prime subunit	NP_290619.1
cyoE	2.11	4.31	0.027937	protoheme IX farnesyltransferase (haeme O biosynthesis)	NP_286170.1
hscA	2.32	4.99	0.012973	heat shock protein, chaperone, member of Hsp70 protein family	NP_289083.1
fimbsub 1	2.53	5.79	0.029658	putative major fimbrial subunit	NP_287648.1

response, which significantly promotes Stx*2* toxin production compared to LII isolates. We have also considered that LII isolates may have mutations affecting the integrity of the *stx2* prophage's late regulatory transcripts shown to encode *stx2* [28-31]. Future work looking at the actual Stx*2* toxin levels as well as evaluation of the structural integrity of the Stx2 phage in these 6 isolates via sequencing or PCR would be a beneficial follow up to this

research. We have performed Stx2b specific ELISA and quantitative PCR analysis of 20 additional LI and 20 additional LII isolates as part of a follow-up study, and found that the LI isolates have statistically ($p < 0.05$) higher transcription rates and protein concentrations under these same conditions (data not shown). If these LII isolates have a defective toxin production system this could be a strong indication that they lack one of the key virulence factors contributing to the pathogenicity of O157 [15,16,32-34].

Urease

Enterohemorrhagic *E. coli* has been shown to be highly adaptable to various extreme environments (water, heat, freezing, acid, desiccation, hypo- and hyperosmotic, disinfectants etc) which contributes greatly to its success as a pathogen [35-46]. To succeed as a enteric pathogen with a low infectious dose [47-49], *E. coli* O157 must be able to survive passage through the acidic environment of the stomach if they are to cause gastrointestinal disease [50]. As an indication of their evolutionary focused ability for surviving acidic environments they possess 3 acid resistance pathways [51] and urease could act as an additional system to modify anion concentrations. Therefore the up-regulation of urease in LI isolates is of interest in spite of recent work indicating that *E. coli* O157 has only rarely been shown to exhibit urease activity [52-55]. As an example, a previous study noted that lack of urease activity in EHEC strains is often due to a base substitution in the *ureD* gene causing an early termination of the transcript [54]. Urease expression and activity be condition, host, or environment specific and could be expressed only in specific environments to beneficially modify internal and/or surrounding anion concentrations, enabling EHEC to survive acidic conditions and contributing to its low infectious dose. Thus, environmental (bovine) isolates may not possess or have sufficient selective pressure for maintenance of detectable levels of urease transcript expression under the conditions evaluated.

Previous research by Heimer et al [52] suggests regulation of the urease operon is through *fur* (not differentially regulated) and an unknown trans-acting factor. It was hypothesized that this transacting factor is missing in *E. coli* O157:H7 strain EDL933 (atcc # 43895) though other O157 strains (IN1 and MO28) have been shown to possess some urease activity. However, none of the isolates showed differential regulation of *fur* which may be an indication that the LI isolates may be differentially expressing this proposed transacting factor, which is promoting up-regulation of the urease operons under the current growth conditions. It is likely that based upon previous evaluation that there is some low level urease activity that is not evident in *E. coli* O157 strains using conventional methods such as Christensen agar [56]. We have begun investigations of the effects of pH, different laboratory media, anoxia, nickel supplementation, and cytosolic specific urease based acid resistance assays on the ability to detect urease activity in O157 isolates.

Curli

Several factors related to attachment are up-regulated in LI isolates. These include curli fibers, type III secretion apparatus genes. This suggests that LI isolates have constitutive up-regulation of many genes that are involved in intimate attachment. It was reported that curli fibers are infrequently expressed during in vitro growth of *E. coli* O157:H7 [8] and that strains containing variations at the *csgD* promoter region, which induced expression of curli, are associated with increased virulence in mice and increased invasion of HEp-2 cells [57]. In this experiment there was significant up-regulation of *csgA* and *csgD* as well as some evidence for up-regulation of the both *csg* operons [see Additional file 1] in the LI isolates compared to LII, yet genes involved in regulation of curli operons do not correspond to this observation. *RpoS* has been shown to interact with *hns* (neither differentially regulated) to derepress *csgAB* expression [58]. Further contradicting the increased expression of curli operons in LI, *ompR* is up-regulated in LII. Increased *ompR* expression has also been associated with increased curli production yet a single point mutation, in *ompR* [59]. Future work should likely evaluate whether curli fibres are actually being produced and assembled under these *in vitro* conditions in LI isolates.

Virulence gene regulation

One of the more interesting of the up-regulated genes in LI is *rfaH*. Originally, discovered as a primary regulator of LPS-core synthesis in *Salmonella enterica* and *E. coli* [60,61], *RfaH* is noted as a primary virulence regulator of *E coli* that functions as a transcriptional anti-terminator [62,63] in long operons. These operons include those encoding the F-factor, O-antigens, different capsules, hemin uptake receptor, alpha-hemolysin, and CNF-1 [64-73]. Inactivation of *rfaH* in uropathogenic *E. coli* has be shown to inhibit pathogenicity completely [74]. *RfaH* mutants have been shown to have reduced ability to survive/grow in the presence of bile salts [75]. The up-regulation of *rfaH* in LI isolates may be an important avenue to pursue as a means to explain their hypothesized enhanced virulence.

LEE

LII isolates showed an increased expression of *toxB* which is known to promote expression of genes encoded by locus of enterocyte effacement (LEE). Indeed, several of the *esp* (*A*, *B*, *P*) showed slight cumulative up-regulation. In addition, most of the *etp* genes involved in the type II transport system were also up-regulated. The type II secre-

tion system was recently noted as also being involved in intimate attachment through secretion of *stcE* [76]. These results showing upregulation of such an important virulence factor in LII isolates points out two key features that are of importance in this manuscript. The first is that these results as intended can help with identification of isolates which may serve as good regulatory models for providing additional insight into virulence expression. In addition, these results are obviously counter to the overall hypothesis that LI is either more virulent or has more potential for transmission and therefore serve as a caution for the interpretation of results. Thus, as with all microarray studies care must be taken in interpretation of the results, yet negative results or results counter to the hypothesis should not be ignored.

LPS, fimbria, and Flagella

LII isolates also show notable up-regulation of genes involved in a number of systems that are noted as virulence factors. Of interest in LII is the comparative up-regulation of LPS, fimbria (FimH), capsule, and flagella related genes (Table 1 and supplement). Considering that the isolates were grown under anaerobic conditions the increase in LPS and flagella related transcripts represents what may be a typical K-12 like *E. coli* response to anoxic conditions [77] in the LII isolates, while the LI seem to be lacking this common profile. The hypothesized decrease virulence of LII may be partially explained by the more pronounced regulation of certain virulence factors by LI. Another interesting aspect that is related to the expression of genes associated with motility and the results seen here is the hypothesis proposed by Monday et al. [78], which is related to a competitive interaction between different type III secretion systems. According to this hypothesis there could be a competitive interaction between the type III secretion systems associated with flagellar export and assembly and the type III secretion system that mediates the injection of virulence factors (LEE). Thus, because O157 has multiple type III systems there is the potential for these systems to interfere with one another. This competition could ultimately affect the expression of motility and/or virulence factors. Thus, because there is an increase in LEE expression as well as motility genes in LII isolates it may be a result of an interaction of the type III regulatory networks in these isolates.

In proper proportions type 1 fimbriae and the LPS of uropathogenic *E. coli* are known to operate together to induce apoptosis in human neutrophils [79]. The cooperative effects of these virulence attributes may function as a mechanism by which *E. coli* induces infections of the urinary tract. However, if LPS is over produced, excess LPS is likely to be secreted by bacteria into their environment, which may have the opposite effect. In fact, it has been documented that if significant amounts of LPS is released from non-adherent bacteria this has an anti-apoptotic effect on neutrophils, suggesting that LPS can also serve as an important regulator of neutrophil survival in tissue [79]. Up regulation of LPS by LII isolates compared to LI isolates, if this excess LPS were shed from the bacteria, maintained in the cytoplasm, or deposited in excess onto the membrane might also be toxic to the bacteria inhibiting its own growth and interaction with its environment [80]. Overproduction of LPS could also alter bacterial cell morphology by accumulation in the bacterial cytosol, which could also potentially prevent pathogenesis. Previous work [81] and [82] demonstrated that *E. coli* O157 exhibiting reduced production of O157 LPS side chains displayed an increased binding to tissue culture cells. It was concluded that the presence of the O157 polysaccharide has the potential to interfere with the adherence and its expression is not required to produce the attaching and effacing lesions. Excess LPS may act to mask adhesive structures present on the bacterial surface. It is also possible that the physicochemical properties of the cell such as surface charge or hydrophobicity may be altered by lack of or excess LPS. These hypothetical interactions and the effects of LPS expression on pathogenesis are again a highly interesting topic for future research.

Conclusion

It has been hypothesized by various researchers that a less pathogenic lineage of *E. coli* O157 exists. Geared toward finding evidence that might direct research toward genetic mechanisms that support the hypothesis of differential virulence or transmission potential we evaluated representatives from these two lineages in a preliminary study. The results highlight several of the more important virulence factors as being differentially regulated, as well as various regulatory networks that may provide useful insight and targets for future research. Key virulence factors were shown to be upregulated in LI, especially those that have been suggested to promote virulence and transmission potential. However, other contradictory findings were also uncovered in which several virulence factors more associated with colonization and pathogenesis were also upregulated in LII isolates. Many previous studies describing regulatory mechanisms are supported by the results of this study, providing some additional insight into the control of virulence genes. Though the hypotheses considered as part of this research is still far from conclusive, the results do provide a valuable foundation that will direct future research. Ultimately, evaluating the expression of key virulence factors among different *E. coli* O157 isolates is valuable beyond the reasoning discussed within the confines of this report, and the interpretation of such expression data will continue as the understanding of virulence, pathogenesis and transmission evolves.

Table 3: Gene Ontology Biological Processes and Molecular Functions representations for level 4 and level 6, presenting genes from the above tables and supplemental data table grouped into those gene ontology categories shown to be statistically over represented.

Gene Ontology Classification	% of total upregulated genes in LI	% of total upregulated genes in LII	p-value
MOLECULAR FUNCTION LEVEL 4			
Peptidase activity	0	16.33	0.0059
Transferase activity, phosphorus transfer	6.67	22.45	0.043
DNA binding	42.22	22.45	0.048
MOLECULAR FUNCTION LEVEL 6			
Urease activity	12.5	0	0.05
GTP binding	12.5	0	0.05
BIOLOGICAL PROCESS LEVEL 4			
Regulation of cellular physiological processes	24.62	8	0.009
Regulation of metabolism	24.62	8	0.009
Nitrogen compound metabolism	6.15	0	0.04
BIOLOGICAL PROCESS LEVEL 6			
Regulation of nucleobase, nucleoside binding	30.61	11.32	0.02
Regulation of transcription	30.6	15.1	0.026

the file format, the title of the data, and a short description of the data

All cells have stress response pathways that help to maintain homeostasis, however it appears that these two lineages of O157 may have diverged just enough that their regulatory pathways are geared for different purposes, ultimately promoting survival in different environments and hosts. It is not clear yet, though research is ongoing, whether LII isolates have lower transmission potential or lower virulence or indeed whether there is enough divergence between the two lineages to consider them as separate. One hypothesis presented in the literature and also supported by the data presented is that LII strains may be more co-evolved as a symbiont of cattle, which promotes its long-term survival in this specific reservoir. For instance, stx2 expression may not be as beneficial in colonization of a bovine host as it has been noted that intestinal receptors for Shiga toxin are found in humans but not cattle [83], while LEE island expression may be very important [84]. Popular theories of pathogen evolution suggest that as a pathogen evolves within finite populations, the pathogen tends to become less virulent (attenuation) to the host thereby promoting though various mechanisms of evolution its own transmission and survival among the populations [85]. This may be exemplified by the differential expression of stress response genes, which could prime or maintain an isolate of *E. coli* O157 in a genetic state that is able to rapidly respond to conditions the isolate might encounter during transmission from animal to human hosts, through the farm to fork process, thereby increasing its transmission potential.

Methods

Bacterial isolates and growth conditions

A working set of lineage (20 LI and 20 LII) isolates as described in Kim et al. [1] were obtained from A. Benson (University of Nebraska, Dept. Of Food Science and Technology, Lincoln, Neb.). LI isolates 43895, fda518, frik533 and LII isolates ne037, frik2000, frik1985 were chosen at random and utilized in the current analyses. Isolates were grown on LB agar under anaerobic conditions for 12 hours. Previous growth studies noted that these 6 isolates displayed similar growth curves, OD600, and concentration (data not shown). Stationary phase was selected to ensure that all isolates and cultures were at the same stage of growth. Isolates were of the Stx2vha genotype and all exhibited typical O157 phenotype characteristics including acid tolerance, lack of sorbitol fermentation, lack of glucuronidase activity and beta hemolysis on tryptose blood agar (Difco, Sparks, MD) with washed, defibrinated sheep blood (Oxoid, Lenexa, KS). All isolates also displayed the same phenotypes using API20 (bioMerieux, Durham, NC).

Microarray design

Using the transcriptome of *E. coli* O157:H7 EDL933 an oligonucleotide microarray (~50mer) was designed. Based upon funding available we were able to choose 610 genes [see Additional file 2] including 10 negative control genes derived from pig sequences, which were selected based upon their being associated with virulence or with regulation of virulence genes. Specifications of oligos were based upon various design characteristics such as temperature of melting, 3' location, specificity, lack of repeat nucleotides, etc. [86]. Oligos were synthesized and normalized in concentration by Integrated DNA Technologies Inc. (Coralville, IA). Oligos were resuspended in Epoxide Slide Spotting Solution and printed onto Epoxide Coated Slides (Corning Inc., Corning, NY). Each array consisted of duplicate elements and each slide contained a duplicate array.

Microarray protocol

All procedures were performed according to respective manufacturer protocols. Colonies were resuspended immediately in RNAprotect Bacteria Reagent (Qiagen Inc., Valencia, CA) after they were harvested. Total RNA was extracted using RNeasy Protect Bacteria Mini Kit (Qiagen Inc.) and DNA removed using RNase-Free DNase Set (Qiagen Inc.). RNA was quantified using a nanodrop ND-1000 device (NanoDrop Technologies, Wilmington, DE) and quality confirmed by electrophoresis. RNA was labeled with either CyDye3-dCTP or CyDye5-dCTP (Amersham Biosciences) using the LabelStar kit (Qiagen Inc.) and Random nonamers (Sigma-Aldrich Inc., St. Louis, MO). Labeled cDNA was hybridized to the microarray using Universal Hybridization Solution (Corning Inc.).

Microarray analysis

Each microarray experiment was performed in duplicate and each experiment also had a corresponding dye swap for an added technical replication. As an example of a dye swap design LI is labeled with cy3 and LII is labeled with cy5 in one array and in the second array LI is labeled with cy5 and LII labeled with cy3. Dye swaps are not biological replicates but provide technical replication that accounts for different dye incorporation rates. Images were captured using a Genepix 4000B (Molecular Devices Corporation, Union City, CA) laser scanner and images processed using GenePix 6.0 software (Molecular Devices Corporation). Analysis was performed using Acuity 4.0 software as well as GeneSpring 11.0 software (Agilent Technologies, Palo Alto, CA). Results were compared between the two software packages to assure conformity of results. Slides were normalized using standard settings (ratio based so that the mean of the ratio of means, of all features, were equal to 1.0). All ratios less than 0.1 and greater than 10.0 were excluded, as well as bad, low signal, absent, or unfound features. To obtain our final data provided in Table 1 and Table 2 we required that all arrays, duplicate elements on each array, and these same features on the dye-swap experiments (after mathematical conversion x' = -x) to provide agreement, show significant relevance at the $p < 0.05$ level, and exhibit at least 1.2 fold regulation. A supplemental dataset was derived for those genes that showed a tendency to be differentially expressed. Usually, the lack of inclusion into the stringent dataset was only based upon the quality of the signal in one of the array or dye swap comparisons. Therefore, these results are provided for information and discussion purposes.

Quantitative PCR

The results of the array were validated using quantitative PCR. Subsets of the regulated genes were chosen at random and primers designed using Primer Select 2.0 software (Applied Biosystems, Foster City, CA). RNA was quantified using NanoDrop system and then using QuantiTect SYBR Green RT-PCR kit (Qiagen Inc.) relative CT was determined with 16s as a control gene, using ABI 7500 Real Time-PCR system (Applied Biosystems).

Functional analysis

HT-GO-FAT software was used to perform the functional GO related analysis. Functional classifications were determined for the regulated genes using HT-GO-FAT and the LIRU8 database. Statistics for higher represented classifications were also determined using HT-GO-FAT. A dedicated Amigo database was also prepared based upon the microarray and the EDL933 transcriptome and can be found at the above URL.

Statistical analysis

Acuity 4.0 built in statistics algorithms were utilized for all statistics related to microarrays. One sample t test was used to determine the significantly regulated genes. Random samples assigned by computer generation. Standard methods were utilized for evaluation of quantitative PCR based upon target gene Ct values (number of cycles of PCR before a threshold of detection is crossed) normalized with the Ct value of an appropriate housekeeping gene (fadD) to compensate for variation in initial RNA and cDNA concentrations. The first normalization procedure provides the initial ΔCt value. The sample ΔCt values were then normalized against the smallest ΔCt value identified in the complete data set, termed ΔΔCt. Finally, the ΔΔCt value for each sample was transformed by the function $2^{\Delta\Delta CT}$ to produce the final gene expression value for each sample. This method allowed for direct comparison of relative gene expression values between isolates. Gene Ontology related statistics were calculated as described by Al-Shahrour et al [87].

Authors' contributions

SD designed the microarray, conceived of the project and wrote the manuscript, HI performed the laboratory experiments.

USDA disclaimer

The use of trade, firm, or corporation names in this publication is for the information and convenience of the reader. Such use does not constitute an official endorsement or approval by the United States Department of Agriculture

Additional material

Additional File 1

Regulated Genes p < 0.2 and > 0.05, This file and dataset contains 105 genes that failed a criteria for inclusion in the primary dataset and their significance test was not less the p = 0.05.

Click here for file

[http://www.biomedcentral.com/content/supplementary/1471-2180-6-30-S1.xls]

Additional File 2

All genes contained in the Escherichia coli O157:H7 virulence array, This file provides a list of all of the genes contained in the 610 gene virulence gene O157:H7 array.

Click here for file

[http://www.biomedcentral.com/content/supplementary/1471-2180-6-30-S2.xls]

References

1. Kim J, Nietfeldt J, Benson AK: **Octamer-based genome scanning distinguishes a unique subpopulation of Escherichia coli O157:H7 strains in cattle.** *Proc Natl Acad Sci U S A* 1999, **96:**13288-13293.
2. Pradel N, Boukhors K, Bertin Y, Forestier C, Martin C, Livrelli V: **Heterogeneity of Shiga toxin-producing Escherichia coli strains isolated from hemolytic-uremic syndrome patients, cattle, and food samples in central France.** *Appl Environ Microbiol* 2001, **67:**2460-2468.
3. Yang Z, Kovar J, Kim J, Nietfeldt J, Smith DR, Moxley RA, Olson ME, Fey PD, Benson AK: **Identification of common subpopulations of non-sorbitol-fermenting, beta-glucuronidase-negative Escherichia coli O157:H7 from bovine production environments and human clinical samples.** *Appl Environ Microbiol* 2004, **70:**6846-6854.
4. Barkocy-Gallagher GA, Arthur TM, Siragusa GR, Keen JE, Elder RO, Laegreid WW, Koohmaraie M: **Genotypic analyses of Escherichia coli O157:H7 and O157 nonmotile isolates recovered from beef cattle and carcasses at processing plants in the Midwestern states of the United States.** *Appl Environ Microbiol* 2001, **67:**3810-3818.
5. McNally A, Roe AJ, Simpson S, Thomson-Carter FM, Hoey DE, Currie C, Chakraborty T, Smith DG, Gally DL: **Differences in levels of secreted locus of enterocyte effacement proteins between human disease-associated and bovine Escherichia coli O157.** *Infect Immun* 2001, **69:**5107-5114.
6. Ritchie JM, Wagner PL, Acheson DW, Waldor MK: **Comparison of Shiga toxin production by hemolytic-uremic syndrome-associated and bovine-associated Shiga toxin-producing Escherichia coli isolates.** *Appl Environ Microbiol* 2003, **69:**1059-1066.
7. M. O, N. B: **Weighing Incentives for Food Safety in Meat and Poultry.** *Amber Waves* 2003, **1:**35-42 [http://www.ers.usda.gov/Amberwaves/April03/Features/WeighingIncentives.htm]. USDA ERS
8. Uhlich GA, Keen JE, Elder RO: **Mutations in the csgD promoter associated with variations in curli expression in certain strains of Escherichia coli O157:H7.** *Appl Environ Microbiol* 2001, **67:**2367-2370.
9. Ryu JH, Kim H, Frank JF, Beuchat LR: **Attachment and biofilm formation on stainless steel by Escherichia coli O157:H7 as affected by curli production.** *Lett Appl Microbiol* 2004, **39:**359-362.
10. Ryu JH, Beuchat LR: **Biofilm formation by Escherichia coli O157:H7 on stainless steel: effect of exopolysaccharide and Curli production on its resistance to chlorine.** *Appl Environ Microbiol* 2005, **71:**247-254.
11. Kim SH, Kim YH: **Escherichia coli O157:H7 adherence to HEp-2 cells is implicated with curli expression and outer membrane integrity.** *J Vet Sci* 2004, **5:**119-124.
12. Cookson AL, Cooley WA, Woodward MJ: **The role of type 1 and curli fimbriae of Shiga toxin-producing Escherichia coli in adherence to abiotic surfaces.** *Int J Med Microbiol* 2002, **292:**195-205.
13. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G: **Gene ontology: tool for the unification of biology. The Gene Ontology Consortium.** *Nat Genet* 2000, **25:**25-29.
14. Werber D, Fruth A, Buchholz U, Prager R, Kramer MH, Ammon A, Tschape H: **Strong association between shiga toxin-producing Escherichia coli O157 and virulence genes stx2 and eae as possible explanation for predominance of serogroup O157 in patients with haemolytic uraemic syndrome.** *Eur J Clin Microbiol Infect Dis* 2003, **22:**726-730.
15. Ritchie JM, Thorpe CM, Rogers AB, Waldor MK: **Critical roles for stx2, eae, and tir in enterohemorrhagic Escherichia coli-induced diarrhea and intestinal inflammation in infant rabbits.** *Infect Immun* 2003, **71:**7129-7139.
16. Ludwig K, Sarkim V, Bitzan M, Karmali MA, Bobrowski C, Ruder H, Laufs R, Sobottka I, Petric M, Karch H, Muller-Wiefel DE: **Shiga toxin-producing Escherichia coli infection and antibodies against Stx2 and Stx1 in household contacts of children with enteropathic hemolytic-uremic syndrome.** *J Clin Microbiol* 2002, **40:**1773-1782.
17. Kimura T, Tani S, Motoki M, Matsumoto Y: **Role of Shiga toxin 2 (Stx2)-binding protein, human serum amyloid P component (HuSAP), in Shiga toxin-producing Escherichia coli infections: assumption from in vitro and in vivo study using HuSAP and anti-Stx2 humanized monoclonal antibody TMA-15.** *Biochem Biophys Res Commun* 2003, **305:**1057-1060.
18. Bonnet R, Souweine B, Gauthier G, Rich C, Livrelli V, Sirot J, Joly B, Forestier C: **Non-O157:H7 Stx2-producing Escherichia coli strains associated with sporadic cases of hemolytic-uremic syndrome in adults.** *J Clin Microbiol* 1998, **36:**1777-1780.
19. Fuchs S, Muhldorfer I, Donohue-Rolfe A, Kerenyi M, Emody L, Alexiev R, Nenkov P, Hacker J: **Influence of RecA on in vivo virulence and Shiga toxin 2 production in Escherichia coli pathogens.** *Microb Pathog* 1999, **27:**13-23.
20. Lusetti SL, Drees JC, Stohl EA, Seifert HS, Cox MM: **The DinI and RecX proteins are competing modulators of RecA function.** *J Biol Chem* 2004, **279:**55073-55079.
21. Pages V, Koffel-Schwartz N, Fuchs RP: **recX, a new SOS gene that is co-transcribed with the recA gene in Escherichia coli.** *DNA Repair (Amst)* 2003, **2:**273-284.
22. Shen G, Xu C, Hu R, Jain MR, Nair S, Lin W, Yang CS, Chan JY, Kong AN: **Comparison of (-)-Epigallocatechin-3-Gallate Elicited Liver and Small Intestine Gene Expression Profiles Between C57BL/6J Mice and C57BL/6J/Nrf2 (-/-) Mice.** *Pharm Res* 2005.
23. Bagdasarian M, Bailone A, Angulo JF, Scholz P, Bagdasarian M, Devoret R: **PsiB, and anti-SOS protein, is transiently expressed by the F sex factor during its transmission to an Escherichia coli K-12 recipient.** *Mol Microbiol* 1992, **6:**885-893.
24. Yasuda T, Morimatsu K, Kato R, Usukura J, Takahashi M, Ohmori H: **Physical interactions between DinI and RecA nucleoprotein filament for the regulation of SOS mutagenesis.** *EMBO J* 2001, **20:**1192-1202.
25. Tyler JS, Mills MJ, Friedman DI: **The operator and early promoter region of the Shiga toxin type 2-encoding bacteriophage 933W and control of toxin expression.** *J Bacteriol* 2004, **186:**7670-7679.
26. Plunkett GIII, Rose DJ, Durfee TJ, Blattner FR: **Sequence of Shiga toxin 2 phage 933W from Escherichia coli O157:H7: Shiga toxin as a phage late-gene product.** *J Bacteriol* 1999, **181:**1767-1778.
27. Koudelka AP, Hufnagel LA, Koudelka GB: **Purification and characterization of the repressor of the shiga toxin-encoding bacteriophage 933W: DNA binding, gene regulation, and autocleavage.** *J Bacteriol* 2004, **186:**7659-7669.
28. Muniesa M, Blanco JE, De SM, Serra-Moreno R, Blanch AR, Jofre J: **Diversity of stx2 converting bacteriophages induced from Shiga-toxin-producing Escherichia coli strains isolated from cattle.** *Microbiology* 2004, **150:**2959-2971.
29. Miyamoto H, Nakai W, Yajima N, Fujibayashi A, Higuchi T, Sato K, Matsushiro A: **Sequence analysis of Stx2-converting phage**

VT2-Sa shows a great divergence in early regulation and replication regions. *DNA Res* 1999, **6:**235-240.
30. Iyoda S, Tamura K, Itoh K, Izumiya H, Ueno N, Nagata K, Togo M, Terajima J, Watanabe H: **Inducible stx2 phages are lysogenized in the enteroaggregative and other phenotypic Escherichia coli O86:HNM isolated from patients.** *FEMS Microbiol Lett* 2000, **191:**7-10.
31. Blanch AR, Garcia-Aljaro C, Muniesa M, Jofre J: **Detection, enumeration and isolation of strains carrying the stx2 gene from urban sewage.** *Water Sci Technol* 2003, **47:**109-116.
32. Mukherjee J, Chios K, Fishwild D, Hudson D, O'Donnell S, Rich SM, Donohue-Rolfe A, Tzipori S: **Human Stx2-specific monoclonal antibodies prevent systemic complications of Escherichia coli O157:H7 infection.** *Infect Immun* 2002, **70:**612-619.
33. Fraser ME, Fujinaga M, Cherney MM, Melton-Celsa AR, Twiddy EM, O'Brien AD, James MN: **Structure of shiga toxin type 2 (Stx2) from Escherichia coli O157:H7.** *J Biol Chem* 2004, **279:**27511-27517.
34. Donohue-Rolfe A, Kondova I, Oswald S, Hutto D, Tzipori S: **Escherichia coli O157:H7 strains that express Shiga toxin (Stx) 2 alone are more neurotropic for gnotobiotic piglets than are isotypes producing only Stx1 or both Stx1 and Stx2.** *J Infect Dis* 2000, **181:**1825-1829.
35. Yang SE, Chou CC: **Growth and survival of Escherichia coli O157:H7 and Listeria monocytogenes in egg products held at different temperatures.** *J Food Prot* 2000, **63:**907-911.
36. Williams RC, Sumner SS, Golden DA: **Survival of Escherichia coli O157:H7 and Salmonella in apple cider and orange juice as affected by ozone and treatment temperature.** *J Food Prot* 2004, **67:**2381-2386.
37. Whiting RC, Golden MH: **Variation among Escherichia coli O157:H7 strains relative to their growth, survival, thermal inactivation, and toxin production in broth.** *Int J Food Microbiol* 2002, **75:**127-133.
38. Uyttendaele M, Jozwik E, Tutenel A, De ZL, Uradzinski J, Pierard D, Debevere J: **Effect of acid resistance of Escherichia coli O157:H7 on efficacy of buffered lactic acid to decontaminate chilled beef tissue and effect of modified atmosphere packaging on survival of Escherichia coli O157:H7 on red meat.** *J Food Prot* 2001, **64:**1661-1666.
39. Uyttendaele M, Taverniers I, Debevere J: **Effect of stress induced by suboptimal growth factors on survival of Escherichia coli O157:H7.** *Int J Food Microbiol* 2001, **66:**31-37.
40. Sage JR, Ingham SC: **Survival of Escherichia coli O157:H7 after freezing and thawing in ground beef patties.** *J Food Prot* 1998, **61:**1181-1183.
41. Ryu JH, Beuchat LR: **Influence of acid tolerance responses on survival, growth, and thermal cross-protection of Escherichia coli O157:H7 in acidified media and fruit juices.** *Int J Food Microbiol* 1998, **45:**185-193.
42. Riordan DC, Duffy G, Sheridan JJ, Whiting RC, Blair IS, McDowell DA: **Effects of acid adaptation, product pH, and heating on survival of Escherichia coli O157:H7 in pepperoni.** *Appl Environ Microbiol* 2000, **66:**1726-1729.
43. Nicholson FA, Groves SJ, Chambers BJ: **Pathogen survival during livestock manure storage and following land application.** *Bioresour Technol* 2005, **96:**135-143.
44. McGee P, Bolton DJ, Sheridan JJ, Earley B, Kelly G, Leonard N: **Survival of Escherichia coli O157:H7 in farm water: its role as a vector in the transmission of the organism within herds.** *J Appl Microbiol* 2002, **93:**706-713.
45. McClure PJ, Hall S: **Survival of Escherichia coli in foods.** *Symp Ser Soc Appl Microbiol* 2000:61S-70S.
46. Maule A: **Survival of verocytotoxigenic Escherichia coli O157 in soil, water and on surfaces.** *Symp Ser Soc Appl Microbiol* 2000:71S-78S.
47. Tuttle J, Gomez T, Doyle MP, Wells JG, Zhao T, Tauxe RV, Griffin PM: **Lessons from a large outbreak of Escherichia coli O157:H7 infections: insights into the infectious dose and method of widespread contamination of hamburger patties.** *Epidemiol Infect* 1999, **122:**185-192.
48. Cornick NA, Helgerson AF: **Transmission and infectious dose of Escherichia coli O157:H7 in swine.** *Appl Environ Microbiol* 2004, **70:**5331-5335.
49. Besser TE, Richards BL, Rice DH, Hancock DD: **Escherichia coli O157:H7 infection of calves: infectious dose and direct contact transmission.** *Epidemiol Infect* 2001, **127:**555-560.
50. Lin J, Smith MP, Chapin KC, Baik HS, Bennett GN, Foster JW: **Mechanisms of acid resistance in enterohemorrhagic Escherichia coli.** *Appl Environ Microbiol* 1996, **62:**3094-3100.
51. Price SB, Wright JC, DeGraves FJ, Castanie-Cornet MP, Foster JW: **Acid resistance systems required for survival of Escherichia coli O157:H7 in the bovine gastrointestinal tract and in apple cider are different.** *Appl Environ Microbiol* 2004, **70:**4792-4799.
52. Heimer SR, Welch RA, Perna NT, Posfai G, Evans PS, Kaper JB, Blattner FR, Mobley HL: **Urease of enterohemorrhagic Escherichia coli: evidence for regulation by fur and a trans-acting factor.** *Infect Immun* 2002, **70:**1027-1031.
53. Nakano M, Iida T, Ohnishi M, Kurokawa K, Takahashi A, Tsukamoto T, Yasunaga T, Hayashi T, Honda T: **Association of the urease gene with enterohemorrhagic Escherichia coli strains irrespective of their serogroups.** *J Clin Microbiol* 2001, **39:**4541-4543.
54. Nakano M, Iida T, Honda T: **Urease activity of enterohaemorrhagic Escherichia coli depends on a specific one-base substitution in ureD.** *Microbiology* 2004, **150:**3483-3489.
55. Friedrich AW, Kock R, Bielaszewska M, Zhang W, Karch H, Mathys W: **Distribution of the urease gene cluster among and urease activities of enterohemorrhagic Escherichia coli O157 isolates from humans.** *J Clin Microbiol* 2005, **43:**546-550.
56. Christensen WB: **Urea decomposition as a means of differentiating Proteus and paracolon cultures from each other and from Salmonella and Shigella types.** *J Bacteriol* 1946, **52:**461-466.
57. Uhlich GA, Keen JE, Elder RO: **Variations in the csgD promoter of Escherichia coli O157:H7 associated with increased virulence in mice and increased invasion of HEp-2 cells.** *Infect Immun* 2002, **70:**395-399.
58. Olsen A, Arnqvist A, Hammar M, Sukupolvi S, Normark S: **The RpoS sigma factor relieves H-NS-mediated transcriptional repression of csgA, the subunit gene of fibronectin-binding curli in Escherichia coli.** *Mol Microbiol* 1993, **7:**523-536.
59. Vidal O, Longin R, Prigent-Combaret C, Dorel C, Hooreman M, Lejeune P: **Isolation of an Escherichia coli K-12 mutant strain able to form biofilms on inert surfaces: involvement of a new ompR allele that increases curli expression.** *J Bacteriol* 1998, **180:**2442-2449.
60. Lindberg AA, Hellerqvist CG: **Rough mutants of Salmonella typhimurium: immunochemical and structural analysis of lipopolysaccharides from rfaH mutants.** *J Gen Microbiol* 1980, **116:**25-32.
61. Creeger ES, Schulte T, Rothfield LI: **Regulation of membrane glycosyltransferases by the sfrB and rfaH genes of Escherichia coli and Salmonella typhimurium.** *J Biol Chem* 1984, **259:**3064-3069.
62. Bailey MJ, Hughes C, Koronakis V: **In vitro recruitment of the RfaH regulatory protein into a specialised transcription complex, directed by the nucleic acid ops element.** *Mol Gen Genet* 2000, **262:**1052-1059.
63. Artsimovitch I, Landick R: **The transcriptional regulator RfaH stimulates RNA chain synthesis after recruitment to elongation complexes by the exposed nontemplate DNA strand.** *Cell* 2002, **109:**193-203.
64. Wang L, Jensen S, Hallman R, Reeves PR: **Expression of the O antigen gene cluster is regulated by RfaH through the JUMPstart sequence.** *FEMS Microbiol Lett* 1998, **165:**201-206.
65. Stevens MP, Hanfling P, Jann B, Jann K, Roberts IS: **Regulation of Escherichia coli K5 capsular polysaccharide expression: evidence for involvement of RfaH in the expression of group II capsules.** *FEMS Microbiol Lett* 1994, **124:**93-98.
66. Sanderson KE, Stocker BA: **Gene rfaH, which affects lipopolysaccharide core structure in Salmonella typhimurium, is required also for expression of F-factor functions.** *J Bacteriol* 1981, **146:**535-541.
67. Nagy G, Dobrindt U, Kupfer M, Emody L, Karch H, Hacker J: **Expression of hemin receptor molecule ChuA is influenced by RfaH in uropathogenic Escherichia coli strain 536.** *Infect Immun* 2001, **69:**1924-1928.
68. Leeds JA, Welch RA: **RfaH enhances elongation of Escherichia coli hlyCABD mRNA.** *J Bacteriol* 1996, **178:**1850-1857.

69. Landraud L, Gibert M, Popoff MR, Boquet P, Gauthier M: **Expression of cnf1 by Escherichia coli J96 involves a large upstream DNA region including the hlyCABD operon, and is regulated by the RfaH protein.** *Mol Microbiol* 2003, **47:**1653-1667.
70. Bailey MJ, Koronakis V, Schmoll T, Hughes C: **Escherichia coli HlyT protein, a transcriptional activator of haemolysin synthesis and secretion, is encoded by the rfaH (sfrB) locus required for expression of sex factor and lipopolysaccharide genes.** *Mol Microbiol* 1992, **6:**1003-1012.
71. Marolda CL, Valvano MA: **Promoter region of the Escherichia coli O7-specific lipopolysaccharide gene cluster: structural and functional characterization of an upstream untranslated mRNA sequence.** *J Bacteriol* 1998, **180:**3070-3079.
72. Rahn A, Whitfield C: **Transcriptional organization and regulation of the Escherichia coli K30 group 1 capsule biosynthesis (cps) gene cluster.** *Mol Microbiol* 2003, **47:**1045-1060.
73. Clarke BR, Pearce R, Roberts IS: **Genetic organization of the Escherichia coli K10 capsule gene cluster: identification and characterization of two conserved regions in group III capsule gene clusters encoding polysaccharide transport functions.** *J Bacteriol* 1999, **181:**2279-2285.
74. Nagy G, Dobrindt U, Schneider G, Khan AS, Hacker J, Emody L: **Loss of regulatory protein RfaH attenuates virulence of uropathogenic Escherichia coli.** *Infect Immun* 2002, **70:**4406-4413.
75. Nagy G, Dobrindt U, Grozdanov L, Hacker J, Emody L: **Transcriptional regulation through RfaH contributes to intestinal colonization by Escherichia coli.** *FEMS Microbiol Lett* 2005, **244:**173-180.
76. Grys TE, Siegel MB, Lathem WW, Welch RA: **The StcE protease contributes to intimate adherence of enterohemorrhagic Escherichia coli O157:H7 to host cells.** *Infect Immun* 2005, **73:**1295-1303.
77. Landini P, Zehnder AJ: **The global regulatory hns gene negatively affects adhesion to solid surfaces by anaerobically grown Escherichia coli by modulating expression of flagellar genes and lipopolysaccharide production.** *J Bacteriol* 2002, **184:**1522-1529.
78. Monday SR, Minnich SA, Feng PC: **A 12-base-pair deletion in the flagellar master control gene flhC causes nonmotility of the pathogenic German sorbitol-fermenting Escherichia coli O157:H- strains.** *J Bacteriol* 2004, **186:**2319-2327.
79. Blomgran R, Zheng L, Stendahl O: **Uropathogenic Escherichia coli triggers oxygen-dependent apoptosis in human neutrophils through the cooperative effect of type 1 fimbriae and lipopolysaccharide.** *Infect Immun* 2004, **72:**4570-4578.
80. Ogura T, Inoue K, Tatsuta T, Suzaki T, Karata K, Young K, Su LH, Fierke CA, Jackman JE, Raetz CR, Coleman J, Tomoyasu T, Matsuzawa H: **Balanced biosynthesis of major membrane components through regulated degradation of the committed enzyme of lipid A biosynthesis by the AAA protease FtsH (HflB) in Escherichia coli.** *Mol Microbiol* 1999, **31:**833-844.
81. Bilge SS, Vary JCJ, Dowell SF, Tarr PI: **Role of the Escherichia coli O157:H7 O side chain in adherence and analysis of an rfb locus.** *Infect Immun* 1996, **64:**4795-4801.
82. Cockerill FIII, Beebakhee G, Soni R, Sherman P: **Polysaccharide side chains are not required for attaching and effacing adhesion of Escherichia coli O157:H7.** *Infect Immun* 1996, **64:**3196-3200.
83. Pruimboom-Brees IM, Morgan TW, Ackermann MR, Nystrom ED, Samuel JE, Cornick NA, Moon HW: **Cattle lack vascular receptors for Escherichia coli O157:H7 Shiga toxins.** *Proc Natl Acad Sci U S A* 2000, **97:**10325-10329.
84. an-Nystrom EA, Bosworth BT, Moon HW, O'Brien AD: **Escherichia coli O157:H7 requires intimin for enteropathogenicity in calves.** *Infect Immun* 1998, **66:**4560-4563.
85. Nowak MA, Sasaki A, Taylor C, Fudenberg D: **Emergence of cooperation and evolutionary stability in finite populations.** *Nature* 2004, **428:**646-650.
86. Charbonnier Y, Gettler B, Francois P, Bento M, Renzoni A, Vaudaux P, Schlegel W, Schrenzel J: **A generic approach for the design of whole-genome oligoarrays, validated for genomotyping, deletion mapping and gene expression analysis on Staphylococcus aureus.** *BMC Genomics* 2005, **6:**95.
87. Al-Shahrour F, az-Uriarte R, Dopazo J: **FatiGO: a web tool for finding significant associations of Gene Ontology terms with groups of genes.** *Bioinformatics* 2004, **20:**578-580.

Additional File 1:
Regulated Genes p < 0.2 and > 0.05, This file and dataset contains 105 genes that failed a criteria for inclusion in the primary dataset and their significance test was not less the p = 0.05.

id	average m	average fo	gene	One Samp	definition	
12758269	-3.652	-12.5708	typeIIIap2	0.171781	type III apparatus proten	NP_289421.1
12757814	-2.266	-4.80988	csgF	0.110121	curli production assembly/transport component, 2nd curli operon	NP_287172.1
12757806	-2.1935	-4.57414	cspH	0.146643	cold shock-like protein	NP_286925.1
12757816	-2.12233	-4.35398	csgA	0.153254	curlin major subunit, coiled surface structures; cryptic	NP_287176.1
12757968	-1.739	-3.33804	fliP	0.148823	flagellar biosynthesis	NP_288409.1
12757697	-1.6305	-3.0962	colicin2	0.093691	putative outer membrane receptor for iron compound or colicin	NP_288504.1
12757815	-1.3272	-2.50915	csgE	0.167484	curli production assembly/transport component, 2nd curli operon	NP_287173.1
12757937	-1.2875	-2.44105	chap4	0.094554		
12757997	-1.253	-2.38337	nrfB	0.096684	formate-dependent nitrite reductase; a penta-haeme cytochrome c	NP_290704.1
12758060	-1.22783	-2.34215	ureD_2	0.056155	putative urease accessory protein D	NP_287085.1
12757750	-1.1675	-2.24622	ybbI	0.149632	putative transcriptional regulator	NP_286233.1
12757859	-1.1595	-2.2338	feoA	0.080061	ferrous iron transport protein A	NP_289948.1
12757883	-1.1185	-2.17121	osmB	0.103254	osmotically inducible lipoprotein	NP_287914.1
12758172	-1.09083	-2.12997	terW_2	0.10558	unknown associated with putative tellurite resistance	NP_287107.1
12758134	-1.09017	-2.12899	regDicB	0.107797		
12757813	-1.085	-2.12138	csgG	0.146969	curli production assembly/transport component, 2nd curli operon	NP_287171.1
12757911	-1.0164	-2.02286	puttrans	0.079851		
12757771	-0.97317	-1.96314	emrD	0.061838	2-module integral membrane pump; multidrug resistance	NP_290312.1
12757837	-0.93717	-1.91476	etpC	0.074376	EtpC	NP_052608.1
12757766	-0.91367	-1.88383	Kdolaur	0.080353	(Kdo)2-(lauroyl)-lipid IVA acyltransferase	NP_052689.1
12757739	-0.895	-1.85961	secret2	0.084733	putative secreted protein	NP_288396.
12758131	-0.8435	-1.7944	stx2b	0.121793	shiga-like toxin II B subunit encoded by bacteriophage BP-933W	NP_286977.1
12757873	-0.83617	-1.7853	marA	0.168723	multiple antibiotic resistance; transcriptional activator of defense systems	NP_287621.1
12757747	-0.835	-1.78386	chacha5	0.117391		
12757872	-0.82	-1.76541	marR	0.148274	multiple antibiotic resistance protein; repressor of mar operon	NP_287622.1
12758068	-0.80867	-1.75159	ureA_2	0.051244	putative urease structural subunit A (gamma)	NP_287086.1
12757871	-0.79767	-1.73829	marB	0.118138	multiple antibiotic resistance protein	NP_287620.1
12758223	-0.79233	-1.73187	dinI	0.112327	putative damage induced protein I	NP_287195.1
12758069	-0.73567	-1.66517	ureB	0.121793	putative urease structural subunit B (beta)	NP_286679.1
12757992	-0.73333	-1.66248	fliT	0.084783	flagellar biosynthesis; repressor of class 3a and 3b operons (RflA activity)	NP_288387.1
12758067	-0.723	-1.65061	ureA	0.159273	putative urease structural subunit A (gamma)	NP_286678.1
12757706	-0.67183	-1.5931	terZ	0.114706	putative phage inhibition, colicin resistance and tellurite resistance protein	NP_286706.1
12757791	-0.65483	-1.57443	srmB	0.184226	ATP-dependent RNA helicase	NP_289134.1
12757798	-0.6266	-1.54392	htpG	0.082716	chaperone Hsp90, heat shock protein C 62.5	NP_286214.1
12757921	-0.612	-1.52838	grpE	0.098748	phage lambda replication; host DNA synthesis; heat shock protein; protein repair	NP_289166.1
12758088	-0.60633	-1.52239	wzzB	0.074677	regulator of length of O-antigen component of lipopolysaccharide chains	NP_288532.1
12757880	-0.59225	-1.5076	ybbJ	0.084585	orf, hypothetical protein	NP_286237.1
12758091	-0.589	-1.5042	repFIB	0.114126	RepFIB	NP_052630.2
12758118	-0.588	-1.50316	sbmA	0.051279	sensitivity to microcin B17, possibly envelop protein	NP_286115.1
12757773	-0.58533	-1.50039	rpsD	0.062847	30S ribosomal subunit protein S4	NP_289857.1

Additional File 1:
Regulated Genes p < 0.2 and > 0.05, This file and dataset contains 105 genes that failed a criteria for inclusion in the primary dataset and their significance test was not less the p = 0.05.

id	average m	average fo	gene	One Samp	definition	
12757774	-0.51333	-1.42734	kdtA	0.143234	3-deoxy-D-manno-octulosonic-acid transferase (KDO transferase)	NP_290213.1
12757919	-0.49883	-1.41307	nupC	0.195239	permease of transport system for 3 nucleosides	NP_288963.1
12757824	-0.49517	-1.40948	recC	0.149085	DNA helicase, ATP-dependent dsDNA/ssDNA exonuclease V subunit, ssDNA endonuclease	NP_289374.1
12758040	-0.47717	-1.39201	entC	0.136866	isochorismate hydroxymutase 2, enterochelin biosynthesis	NP_286320.1
12758173	-0.4652	-1.38051	IS600	0.100737		
12758228	-0.41433	-1.33268	kdtB	0.17553	putative enzyme of LPS biosynthesis	NP_290214.1
12758245	-0.405	-1.32409	htrE	0.161465	putative fimbrial usher protein	NP_285835.1
12757858	-0.3874	-1.30803	hemH	0.104084	ferrochelatase: final enzyme of heme biosynthesis	NP_286216.1
12758072	-0.38067	-1.30194	ureC_2	0.161841	putative urease structural subunit C (alpha)	NP_287088.1
12757869	-0.37733	-1.29894	emrA	0.070928	multidrug resistance secretion protein	NP_289236.1
12757685	-0.3535	-1.27766	ybbM	0.074003	putative metal resistance protein	NP_286240.1
12757796	-0.2865	-1.21968	mrdA	0.164274	cell elongation, e phase; peptidoglycan synthetase; penicillin-binding protein 2	NP_286361.1
12757749	-0.2752	-1.21016	ybaO	0.152069	putative transcriptional regulator	NP_286189.1
12758167	-0.25117	-1.19017	rfe	0.130338	UDP-GlcNAc:undecaprenylphosphate GlcNAc-1-phosphate transferase; synthesis of enterobacterial common antigen (ECA)	NP_290415.1
12758124	-0.21275	-1.1589	barA	0.066626	sensor-regulator, activates OmpR by phophorylation	NP_289340.1
12757829	-0.1055	-1.07587	motB	0.192215	enables flagellar motor rotation, linking torque machinery to cell wall	NP_288326.1
12758129	0.157833	1.11561	stx1B	0.091912	shiga-like toxin 1 subunit B encoded within prophage CP-933V	NP_288672.1
12757903	0.2715	1.207062	pqiA	0.208253	paraquat-inducible protein A	NP_286825.1
12758066	0.354833	1.278838	ureG_2	0.199492	putative urease accessory protein G	NP_287091.1
12757898	0.3922	1.312393	fepA	0.173133	outer membrane receptor for ferric enterobactin (enterochelin) and colicins B and D	NP_286310.1
12757711	0.401333	1.320728	terE	0.063892	putative phage inhibition, colicin resistance and tellurite resistance protein	NP_286711.1
12758211	0.4738	1.388763	ABCtrans2	0.092367		
12758286	0.4905	1.404932	ybgL	0.092461	putative lactam utilization protein	NP_286428.1
12757673	0.520667	1.434618	waaJ	0.173952	putative LPS biosynthesis enzyme	NP_290207.1
12757677	0.531	1.44493	waaY	0.162598	putative LPS biosynthesis protein	NP_290208.1
12757887	0.5498	1.463883	blc	0.143489	outer membrane lipoprotein (lipocalin)	NP_290784.1
12758122	0.565	1.479388	basS	0.125813	sensor protein for basR	NP_290745.1
12758120	0.565167	1.479558	rstB	0.135925	sensor histidine protein kinase (RstA regulator)	NP_288044.1
12757879	0.57	1.484524	wzy	0.192132	O antigen polymerase	NP_288545.1
12757839	0.571	1.485553	etpE	0.101982	EtpE	NP_052610.1
12757689	0.605333	1.52133	ycfX	0.182615	putative NAGC-like transcriptional regulator	NP_287253.1
12758109	0.610167	1.526436	rpoD	0.174049	RNA polymerase, sigma(70) factor; regulation of proteins induced at high temperatures	NP_289642.1
12757710	0.611167	1.527494	terD	0.083244	putative phage inhibition, colicin resistance and tellurite resistance protein	NP_286710.1
12757716	0.613	1.529436	terD_2	0.111324	putative phage inhibition, colicin resistance and tellurite resistance protein	NP_287118.1
12757932	0.61525	1.531823	cheW	0.138808	positive regulator of CheA protein activity	NP_288324.1
12757955	0.6358	1.553799	secD	0.051975	protein secretion; membrane protein, part of the channel	NP_286147.1

Additional File 1:
Regulated Genes p < 0.2 and > 0.05, This file and dataset contains 105 genes that failed a criteria for inclusion in the primary dataset and their significance test was not less the p = 0.05.

id	average m	average fo	gene	One Samp	definition	
12757674	0.642	1.560491	waaI	0.170411	putative LPS biosynthesis enzyme	NP_290209.1
12757884	0.665833	1.586484	osmC	0.167457	osmotically inducible protein	NP_287673.1
12757918	0.708	1.633538	htrA	0.164557	periplasmic serine protease Do; heat shock protein HtrA	NP_285857.1
12758176	0.730667	1.659406	hemX	0.101321	uroporphyrinogen III methylase	NP_290431.1
12757800	0.733	1.662092	cheZ	0.096421	chemotactic response; CheY protein phophatase; antagonist of CheY as switch regulator	NP_288318.1
12757893	0.7375	1.667284	OMP3b	0.174827	outer membrane protein 3b (a), protease VII	NP_287408.1
12757823	0.745	1.675974	recB	0.058257	DNA helicase, ATP-dependent dsDNA/ssDNA exonuclease V subunit, ssDNA endonuclease	NP_289372.1
12757772	0.760833	1.694469	rpsC	0.130579	30S ribosomal subunit protein S3	NP_289875.1
12758111	0.771667	1.707241	rne	0.137527	RNase E, membrane attachment, mRNA turnover, maturation 5S RNA	NP_287218.1
12757783	0.781167	1.71852	arcB	0.085027	aerobic respiration sensor-response protein; histidine protein kinase/phosphatase, sensor for arcA	NP_289784.1
12758102	0.788	1.726679	rpoA	0.138202	RNA polymerase, alpha subunit	NP_289856.1
12757767	0.810167	1.753414	relA	0.092266	(p)ppGpp synthetase I (GTP pyrophosphokinase); regulation of RNA synthesis; stringent factor	NP_289338.1
12758221	0.8128	1.756617	cytotox	0.139461	putative cytotoxin	NP_289557.1
12757690	0.8164	1.761006	yhcI	0.069253	putative NAGC-like transcriptional regulator	NP_289790.1
12758123	0.845167	1.796472	phoQ	0.10352	sensor protein PhoQ	NP_287343.1
12757801	0.85775	1.81221	cheY	0.071958	chemotaxis regulator transmits chemoreceptor signals to flagelllar motor components	NP_288319.1
12757904	0.861667	1.817136	pqiB	0.103552	paraquat-inducible protein B	NP_286826.1
12757789	0.8695	1.82703	glnQ	0.095217	ATP-binding component of glutamine high-affinity transport system	NP_286573.1
12758006	0.875167	1.83422	glnP	0.086466	glutamine high-affinity transport system; membrane component	NP_286574.1
12757946	0.917	1.888185	wecG	0.167265	probable UDP-N-acetyl-D-mannosaminuronic acid transferase; synthesis of enterobacterial common antigen (ECA)	NP_290426.1
12758271	0.918	1.889494	intcoloniz	0.205343	putative intestinal colonization factor	NP_287515.1
12757672	0.95075	1.932877	waaD	0.122001	putative LPS biosynthesis enzyme	NP_290206.1
12758250	1.008	2.011121	fimbrial3	0.148336	fimbrial subunit	
12758222	1.036	2.050534	cytotox2	0.072327	putative cytotoxin	NP_289558.1
12757969	1.0995	2.142804	fliQ	0.063168	flagellar biosynthesis	NP_288410.1
12757701	1.1525	2.222988	oxidoreduc	0.137788	putative oxidoreductase	
12758049	1.2775	2.424185	fimA	0.075612	major type 1 subunit fimbrin (pilin)	NP_290930.1
12757723	1.5065	2.841199	protease1	0.06138	putative protease	NP_285915.1
12757803	1.588833	3.00806	cspD	0.051261	cold shock protein	NP_286652.1

Additional File 2:
All genes contained in the Escherichia coli O157:H7 virulence array, This file provides a list of all of the genes contained in the 610 gene virulence gene O157:H7 array.

id	Name
12757844	etpK
12757846	etpM
12758121	hydH
12758206	ykgA
12757761	ydeW
12757704	fliY
12758236	chaper2
12758156	nupG
12757860	feoB
12758104	rpoB
12757843	etpI
12757999	fumA
12757935	phoB
12758152	rfaH
12758046	argT
12758212	yejH
12757974	flgC
12757845	etpL
12758032	oppC
12757842	etpH
12758240	fimb10
12758019	hscA
12757878	wzx
12758095	cheB
12758000	fumC
12757895	fhuA
12757833	etpN
12757752	yegW
12757832	etpJ
12758016	hslJ
12757810	cutC
12758251	fimberal4
12757988	fliC
12758031	cspG
12757745	trxC
12758258	helicase1
12758094	ompR
12758148	toxB
12758103	rpoC
12757834	etpO
12758144	wecF
12758130	stx2A
12758246	usher2
12757835	espP
12757740	secret3
12757730	ydeA
12757956	cyoE
12757927	inaA
12757680	fimbsub1

Additional File 2:
All genes contained in the Escherichia coli O157:H7 virulence array, This file provides a list of all of the genes contained in the 610 gene virulence gene O157:H7 array.

id	Name
12757936	chap3
12758014	ibpA
12757888	pldA
12757954	secF
12758160	typeIIIap7
12757684	transreg
12757809	YjhSpro
12758013	ibpB
12757722	ybbK
12757759	ycjZ
12758070	ureB_2
12758059	ureD
12758077	virulense3
12758041	kfras
12758001	frdD
12758171	terW
12757780	hemY
12757804	cspC
12758068	ureA_2
12757803	cspD
12758118	sbmA
12757955	secD
12758060	ureD_2
12757823	recB
12757723	protease1
12757771	emrD
12757773	rpsD
12757969	fliQ
12757711	terE
12758124	barA
12757690	yhcI
12757869	emrA
12757801	cheY
12758222	cytotox2
12757685	ybbM
12757837	etpC
12758088	wzzB
12758049	fimA
12757911	puttrans
12757859	feoA
12757766	Kdolaur
12757798	htpG
12757710	terD
12757880	ybbJ
12757739	secret2
12757992	fliT
12757783	arcB
12758006	glnP
12758129	stx1B

Additional File 2:
All genes contained in the Escherichia coli O157:H7 virulence array, This file provides a list of all of the genes contained in the 610 gene virulence gene O157:H7 array.

id	Name
12757767	relA
12758211	ABCtrans2
12758286	ybgL
12757697	colicin2
12757937	chap4
12757789	glnQ
12757800	cheZ
12757997	nrfB
12757921	grpE
12758173	IS600
12758176	hemX
12757839	etpE
12757883	osmB
12758123	phoQ
12757904	pqiB
12757858	hemH
12758172	terW_2
12758134	regDicB
12757814	csgF
12757716	terD_2
12758223	dinI
12758091	repFIB
12757706	terZ
12757747	chacha5
12757871	marB
12758069	ureB
12757672	waaD
12758122	basS
12758167	rfe
12757772	rpsC
12758120	rstB
12758040	entC
12758111	rne
12757701	oxidoreduc
12758102	rpoA
12757932	cheW
12758221	cytotox
12757774	kdtA
12757887	blc
12757806	cspH
12757813	csgG
12757872	marR
12758250	fimbrial3
12757968	fliP
12757824	recC
12757750	ybbI
12757749	ybaO
12757816	csgA
12758067	ureA

Additional File 2:
All genes contained in the Escherichia coli O157:H7 virulence array, This file provides a list of all of the genes contained in the 610 gene virulence gene O157:H7 array.

id	Name
12758245	htrE
12758072	ureC_2
12757677	waaY
12757796	mrdA
12757918	htrA
12757946	wecG
12757884	osmC
12757815	csgE
12757873	marA
12757674	waaI
12758269	typeIIIap2
12757898	fepA
12757673	waaJ
12758109	rpoD
12757893	OMP3b
12758228	kdtB
12757689	ycfX
12757791	srmB
12757879	wzy
12757829	motB
12757919	nupC
12758066	ureG_2
12758271	intcoloniz
12757903	pqiA
12757841	etpG
12758149	tagA
12758220	Memspan
12757698	OMRiron
12758154	slyA
12757712	terZ_2
12758132	rseC
12758131	stx2B
12758166	lpxD
12758036	himA
12758107	rpoH
12758235	fimbchap12
12757792	bacA
12758192	psiB
12758084	flhC
12757943	secA
12757900	cirA
12757769	rsuA
12758089	oraA
12757840	etpF
12757951	secB
12757923	pspD
12757770	entA
12757890	ompF
12758048	pldB

Additional File 2:
All genes contained in the Escherichia coli O157:H7 virulence array, This file provides a list of all of the genes contained in the 610 gene virulence gene O157:H7 array.

id	Name
12758029	oppD
12757707	terA
12757913	def
12757686	fimb5
12758117	espD
12758196	csgD
12758143	chpR
12758204	chacha2
12757822	recD
12757924	pspB
12757920	hipA
12758064	ureF_2
12758242	fimb8
12757938	ccmH
12758098	hydG
12758219	csgC
12758147	dsbD
12758065	ureG
12757831	fes
12758285	ycjW
12757914	mrcA
12757802	cspE
12758262	holin2
12757973	flgB
12757957	motA
12757748	ecnB
12758108	rpoN
12758180	yehE
12758205	ecnA
12757775	rplC
12757795	cls
12757950	secG
12758035	fic
12758203	ybaT
12758105	rpoZ
12758008	ybaY
12757786	ascG
12757870	emrB
12757785	aroP
12758112	mrdB
12757776	rplD
12758279	ironperm3
12757926	pspC
12757929	hemC
12757817	ccmE
12757793	bcp
12758127	degQ
12757848	hcaC
12757945	yhbH

Additional File 2:
All genes contained in the Escherichia coli O157:H7 virulence array, This file provides a list of all of the genes contained in the 610 gene virulence gene O157:H7 array.

id	Name
12758274	ygdP
12758055	methylvio
12758027	hlyC
12758190	control10
12758011	hha
12757928	pmrD
12758110	rpoE
12757713	terA_2
12758229	yhcK
12758080	qseC
12758261	holin1
12758101	RfbU
12758227	enterotox
12758097	glnG
12758281	ironperm1
12757715	terC_2
12757733	yggT
12758119	rcsC
12757944	chpA
12758007	glnS
12757910	terF_2
12758099	rstA
12758137	umuC
12757849	fpr
12758139	parB
12757737	ygeA
12758071	ureC
12758076	virulence2
12758093	repA1
12757861	finO
12757734	yhjX
12758003	rfaG
12757741	putsec4
12758061	ureE
12758028	hlyD
12757768	spoT
12757718	acyltrans
12758043	letA
12758200	yebL
12757972	flgG
12757925	pspA
12758287	dedD
12757808	nusG
12757967	fliL
12758280	ironrecep8
12757825	radC
12758125	cheA
12758004	gadB
12758054	tap

Additional File 2:
All genes contained in the Escherichia coli O157:H7 virulence array, This file provides a list of all of the genes contained in the 610 gene virulence gene O157:H7 array.

id	Name
12758178	hemD
12757885	tolC
12757983	flgA
12758199	adhesin1
12757709	terC
12758260	heminlipo
12757763	transp4
12758175	exbB
12757866	modA
12758052	tar
12758085	flhD
12757708	terB
12757881	fliZ
12757799	dnaJ
12757909	terF
12757917	glnH
12758106	rpoS
12758257	ddg
12758278	ironperm2
12758213	sidI
12758009	mopA
12758151	nusB
12758231	ybgD
12757981	flgD
12758193	rtcR
12758053	trg
12757676	waaQ
12758238	fimb4
12757912	dacD
12758012	clpB
12757894	slp
12758096	cheR
12758145	tehB
12757892	ompA
12758113	mreD
12757671	waaL
12757788	fepC
12758024	ccmC
12757940	ppdB
12757675	waaP
12757838	etpD
12758215	cellkill
12757781	wbdR
12757724	putrecep
12757830	entD
12757953	glnD
12757876	nac
12758140	sspB
12757855	fepG

Additional File 2:
All genes contained in the Escherichia coli O157:H7 virulence array, This file provides a list of all of the genes contained in the 610 gene virulence gene O157:H7 array.

id	Name
12758025	hlyA
12758162	typeIIIap9
12757989	flgL
12757961	fimG
12757875	nikB
12757805	cspA
12758197	eaeH
12758044	letB
12757754	yihL
12757960	fimF
12758057	permeas
12757891	ompC
12758078	qseA
12758218	compresist
12758259	chuS
12757987	fliD
12757821	cutA
12758273	invasin4
12757826	parE
12757930	rcsA
12758142	msbB
12758150	nusA
12757857	yecI
12758226	endolys3
12758174	exbD
12757896	ompX
12757790	clpP
12758075	mviN
12757714	terB_2
12758177	cysG
12758168	wecB
12758136	sopB
12757699	prrA
12758239	fimb9
12758116	espB
12757856	ubiB
12758179	sanA
12758018	hslV
12758133	rseA
12757922	pspE
12757700	ybgQ
12757907	OMPLom
12758255	yciM
12758017	hslU
12757746	putoxin
12758210	adhesin2
12757980	flgK
12757964	fliO
12757868	emrY

Additional File 2:
All genes contained in the Escherichia coli O157:H7 virulence array, This file provides a list of all of the genes contained in the 610 gene virulence gene O157:H7 array.

id	Name
12758045	lysU
12758195	ygaA
12757693	yciD
12757762	tir
12757952	flgN
12757779	ascB
12758224	yihW
12757744	ybbN
12758170	uspA
12758214	ybbB
12757915	mrcB
12758090	glnB
12757738	CyMexp
12758155	sdiA
12757764	ybgR
12758146	tehA
12758135	sopA
12757683	permalt
12758062	ureE_2
12757899	fhuE
12757863	csgB
12757908	OMPLom2
12757847	uvrC
12758282	bactferrin
12757819	hmpA
12757902	ycbF
12757982	fliA
12757725	ybbU
12757934	rcsB
12758083	rseB
12758253	yihK
12757886	chuA
12757947	degS
12757942	secE
12757949	ptr
12757865	modB
12758034	IS629
12757836	ftsN
12758157	tatA
12757991	fliS
12757965	fhiA
12758272	invasin3
12757717	terE_2
12758233	fimbchap8
12757742	citA
12757998	fsr
12757736	ygeD
12757827	hns
12757692	Omexp

Additional File 2:
All genes contained in the Escherichia coli O157:H7 virulence array, This file provides a list of all of the genes contained in the 610 gene virulence gene O157:H7 array.

id	Name
12758284	putkiller3
12757778	rplF
12758194	yedW
12758042	klcA
12758022	ccmB
12758169	wecC
12758115	espA
12758023	ccmD
12758086	mreB
12758079	qseB
12758051	metG
12758005	gadA
12757901	papClik
12758074	mviM
12757691	yegT
12758201	adhesin8
12758010	mopB
12757743	ssbW
12758141	sugE
12758158	typeIIIap3
12757948	ptrB
12757828	lon
12757874	fur
12757670	Lom-lik3
12757751	ybiH
12757852	fepB
12758225	endolys1
12758081	fimB
12758037	himD
12758283	kilW
12758237	sfmF
12757784	flgM
12758047	lysP
12757811	cutF
12757732	resist4
12758020	htpX
12757933	gcvA
12757726	yecG
12758128	stx1A
12758087	icc
12758050	melB
12758092	repA2
12758063	ureF
12757782	asr
12758208	ARAC2
12757979	flgE
12757797	dnaK
12758263	gamW
12758114	mreC

Additional File 2:
All genes contained in the Escherichia coli O157:H7 virulence array, This file provides a list of all of the genes contained in the 610 gene virulence gene O157:H7 array.

id	Name
12757694	yehB
12757812	acrF_2
12758202	adhesinFimH
12758207	ARAC3
12757818	ccmF
12757678	chacha4
12758216	chaper1
12757705	chuU
12758126	citB
12757807	compres1
12758182	control2
12758183	control3
12758186	control6
12758187	control7
12758189	control9
12758038	eae
12757867	emrK
12757854	fepD
12757853	fepE
12758249	fimb-lik
12758241	fimb11
12757688	fimb7
12758234	fimbchap11
12757679	fimbsub13
12757681	fimbsub2
12757682	fimbsub3
12758244	fimbush11
12757897	fimD
12758082	fimE
12757864	fimH
12758243	fimush
12757963	fimZ
12758230	fixX
12757975	flgF
12757971	flgH
12758033	flgI
12757966	flgJ
12757990	flhA
12757995	flhE
12757984	fliE
12757985	fliF
12757978	fliG
12757986	fliH
12757996	fliI
12757993	fliJ
12757994	fliK
12757976	fliM
12757970	fliR
12758021	hemaggl

Additional File 2:
All genes contained in the Escherichia coli O157:H7 virulence array, This file provides a list of all of the genes contained in the 610 gene virulence gene O157:H7 array.

id	Name
12758026	hlyB
12757721	hlyE
12758267	hofB
12758266	hofC
12757931	htgA
12758015	htrC
12757669	liptypeIII
12757695	lomK
12757889	ompG
12758030	oppF
12757719	pilin1
12757939	ppdA
12757941	ppdC
12757958	psiA
12758198	putadhes
12757728	regcelldiv2
12757777	rplE
12758247	sfmA
12757696	sfmD
12758163	typeIIIap10
12758164	typeIIIap11
12758159	typeIIIap4
12758270	typeIIIap5
12758165	typeIIIap6
12758161	typeIIIap8
12758058	typeIIIapp
12758138	umuD
12758073	usher1
12757720	wza
12757755	ybbS
12758265	ybdE
12757756	ybdO
12757757	ybhD
12758254	ycaL
12758248	ycbQ
12757758	ydaK
12757760	yeaT
12758256	yegD
12758232	yfcS
12758275	ygeH
12757882	yihM
12757735	yihN
12757731	ylcD